ITIF

The Information Technology
& Innovation Foundation

Before the
Federal Communications Commission
Washington, D.C. 20554

In the Matter of:	)	
	)	
Framework for Broadband Internet Service	)	**GN Docket No. 10-127**
	)	

REPLY COMMENTS OF INFORMATION TECHNOLOGY AND INNOVATION FOUNDATION

Richard Bennett

Information Technology and Innovation Foundation[1]
1101 K St N.W.
Suite 610
Washington, DC 20005

[1] ITIF is a nonprofit, non-partisan public policy think tank committed to articulating and advancing a pro-productivity, pro-innovation and pro-technology public policy agenda internationally, in Washington and in the states. Through its research, policy proposals, and commentary, ITIF is working to advance and support public policies that boost innovation, e-transformation and productivity.

The Information Technology and Innovation Foundation (ITIF) is pleased to offer the following reply comments regarding a regulatory framework for broadband Internet service. ITIF has long advocated a Third Way approach to Internet regulation in the context of the net neutrality controversy and is deeply engaged in the matter of regulatory frameworks that recognize the Internet's unique characteristics and value to society.[2]

1. Summary

Title II reclassification is a regulatory and legal proposal intended to provide the FCC with the authority to impose "network neutrality" regulations on providers of the broadband networks that form a significant portion of the Internet. Reclassification is therefore only a productive approach if one accepts the premise that network neutrality is not only desirable but an urgent necessity. In fact, network neutrality is subtly but importantly different from "Internet openness;" while openness is a desirable goal, neutrality is not. It is not correct to assume that:

1) "Open" and "Neutral" are interchangeable terms.

2) The packet-switched networks interconnected by the Internet are capable of being operated in a manner that is both open and neutral.

3) Voluntary differentiated pricing of Internet service is not a legitimate means of accommodating heterogeneous application requirements and user expectations.

Filers in support of reclassification, such as the Open Internet Coalition (OIC) and its member organizations (Free Press, Public Knowledge, Media Access Project, et. al.) assume that "open and neutral" is a consistent, concrete, practical, and measurable goal and not merely a naïve and metaphorical description. This is an analytical error that betrays a fundamental misunderstanding of the nature of packet-switching, traffic management on broadband networks, and the means by which multiple types of applications are best supported on a common network infrastructure. It also fails to advance national purposes.

[2] Robert D. Atkinson and Philip J. Weiser, *ITIF: A "Third Way" on Network Neutrality*, (Washington, DC: Information Technology and Innovation Foundation, May 30, 2006), http://www.itif.org/index.php?id=63.

At its heart the motivation for reclassification is the no doubt well-intentioned. However, as currently structured the proposal is likely to be counter-productive.

Certainly, the Internet should no more be exempt from regulation than the federal highway system, the global telephone network, or polluting factories, but the regulatory framework that is imposed on the Internet, including last mile broadband networks, must be appropriate according to the nature of its technology and with respect to the national goals and purposes. This framework should be crafted from the bottom up with the Internet's unique nature and purposes in mind, not simply cut, pasted, and adapted from largely unrelated elements of prior technologies.

Therefore, the reclassification exercise is only productive and appropriate as a legal matter if the desired policy goal is productive and appropriate. As the policy goal remains muddled, contradictory, and imprecise, the reclassification exercise is unnecessary at the present time.

Pending Congressional action clarifying the FCC's role in relation to Internet regulation, the FCC can and should convene an expert advisory panel of Internet firms, academics, think tank analysts, and public interest lobbyists to draft guidelines that represent the degree of consensus that exists today regarding Internet regulation. The consensus framework could then be enforced by the mandated refusal of consenting Internet firms to do business with non-compliant dissenters. This power would effectively isolate non-compliers from the Internet, which would spell doom to their economic prospects and serve as a powerful motivator to comply with consensus guidelines. The consensus guidelines should remain in effect until Congressional action clarifies the FCC's role and establishes a legitimate policy framework for Internet regulation. Such a mechanism makes Title II classification unnecessary.

Table of Contents

1. SUMMARY 2

2. TRAFFIC MANAGEMENT IS THE CENTRAL PROBLEM 4

3. OPENNESS AND NEUTRALITY ARE CONTRADICTORY GOALS 8
4. THERE ARE LIMITS TO BANDWIDTH GROWTH 9
5. CONTENT DELIVERY NETWORKS PRIORITIZE PACKETS 11
6. THERE ARE OTHER WAYS TO OBTAIN AUTHORITY 15
7. CONCLUSION 16

2. Traffic Management is the Central Problem

The central dilemma of Internet operation and therefore of Internet regulation is traffic management. ITIF has published a series of reports on this subject (*Managing Broadband Networks*, *Designed for Change*, and *Going Mobile* in particular) that attempt to put the necessity and the benefit of active traffic management in perspective.[3] Despite the enormous body of analytical work on this problem throughout the engineering and policy communities, traffic management remains a locus of misunderstanding.

The misunderstanding and partial understanding of traffic management was evident throughout the FCC's deliberations on the Comcast matter. Many advocates insisted on a *status quo* in which network operators would have been required to refrain from any form of active traffic management whatever. This naïve point of view continues to be widely held, as is evident in recent commentary on the FCC's Open Internet proceeding. A recent ex-parte letter in that proceeding completely mischaracterizes the IETF DiffServ standard in order to buttress this view:[4]

> …it is nonsensical to portray DiffServ as something that a third-party content provider could pay an ISP to use for paid-prioritization. Either an ISP respects DiffServ flags as outlined by IETF and chosen by the

[3] George Ou, *Managing Broadband Networks: A Policymaker's Guide* (Washington, DC: Information Technology and Innovation Foundation, December 2008), http://www.itif.org/files/Network_Management.pdf; Richard Bennett, *Designed for Change: End-to-End Arguments, Internet Innovation, and the Net Neutrality Debate* (Washington, DC: Information Technology and Innovation Foundation, September 2009), http://www.itif.org/index.php?id=294.; and Richard Bennett, *Going Mobile: Technology and Policy Issues in the Mobile Internet* (Washington, DC: Information Technology and Innovation Foundation, March 2010), http://itif.org/publications/going-mobile-technology-and-policy-issues-mobile-internet.

[4] Derek Turner, Free Press, *Re: Notice of Ex Parte Presentation: GN Docket No. 09-191 (Preserving the Open Internet); WC Docket No. 07-52 (Broadband Industry Practices)*, August 3, 2010.

> application or they do not -- and if they do not, then it isn't DiffServ. By way of analogy, an individual customer cannot pay a restaurant to obey the health code -- they either do or they don't. If an ISP is using DiffServ, but not respecting application flags, then that is not the standard as outlined by the IETF. Similar to how Comcast was improperly using RST packets to block BitTorrent, such a nonstandard use of DiffServ would be entirely new, improper, and not at all in line with that outlined by the IETF.

This brief paragraph contains multiple errors of fact:

1. It presumes, on no basis whatever, that DiffServ forbids operators from charging for Differentiated Services. In fact, the creation of the DiffServ architecture was explicitly driven, in part, by the desire to enable operators to create new billable services. In the plain language of the standard:[5]

   > This document defines an architecture for implementing scalable service differentiation in the Internet. A "Service" defines some significant characteristics of packet transmission in one direction across a set of one or more paths within a network. These characteristics may be specified in quantitative or statistical terms of throughput, delay, jitter, and/or loss, or may otherwise be specified in terms of some relative priority of access to network resources. Service differentiation is desired to accommodate heterogeneous application requirements and user expectations, and **to permit differentiated pricing of Internet service**. [Emphasis added]

   What point would there be in differentiating Internet packet services if every level of service had the same price? Every user would simply mark every packet "highest priority" and that would be the end of it.

2. It confuses the method used by DiffServ to signal desired per hop behavior, identifiers known as "DiffServ codepoints" (DSCP), with an earlier method specified by RFC 795 employing precedence flags.[6] Even in RFC 2474, DSCP

[5] S. Blake et al., "RFC 2475 - An Architecture for Differentiated Services," Internet RFC, December 1998, http://tools.ietf.org/rfc/rfc2475.txt.

[6] J. Postel, "RFC 795 - Service mappings," Internet RFC, September 1981, http://tools.ietf.org/html/rfc795.

identifiers do not necessarily have the global intrinsic meaning the ex-parte assigns to them:[7]

> In the packet forwarding path, differentiated services are realized by mapping the codepoint contained in a field in the IP packet header to a particular forwarding treatment, or per-hop behavior (PHB), at each network node along its path. The codepoints may be chosen from a set of mandatory values defined later in this document, from a set of recommended values to be defined in future documents, or **may have purely local meaning**. PHBs are expected to be implemented by employing a range of queue service and/or queue management disciplines on a network node's output interface queue: for example weighted round-robin (WRR) queue servicing or drop-preference queue management. [Emphasis added.]

Consequently, it incorrectly concludes that DiffServ is nothing more than a rule by which the user tells the network how to behave. In fact, it’s a general framework for bilateral communication between an application and a network operator that is adaptable to a wide range of network technologies and service agreements.

3. It confuses a particular DiffServ *implementation* described in RFC 2474 with the overall *architecture* described in RFC 2475.[8]
4. It confuses the RFC 2474 implementation with more current implementations described in RFC 3168, RFC 3246, and RFC 3260, and also fails to comprehend the related work with Explicit Congestion Notification, Integrated Services, and the ongoing IETF work in the Congestion Exposure Working Group (CONEX.)
5. It compares perfectly standard implementations of DiffServ with the generally non-standard (but expedient) use of the TCP RST flag to reduce traffic volume and judges that the only “proper” way to manage traffic is in accordance with some particular RFCs. In fact, there should be no presumption that any RFC imposes an absolute rule on the behavior of any network operator. Adherence to

[7] K. Nichols et al., “RFC 2474 - Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers,” Internet RFC, December 1998, http://tools.ietf.org/rfc/rfc2474.txt.

[8] The title of RFC 2475 provides the clue: “An Architecture for Differentiated Services.”

> RFCs is strictly voluntary throughout the Internet, and there's always a time lag between implementation of new practices and their memorialization in RFCs. The freedom to innovate is the key that prevents the Internet from sinking into the stagnation that has afflicted the public switched telephone network over the past 50 years despite vigorous oversight. While oversight is necessary, it must remain sufficiently permissive as to allow networks to continue on the arc of improvement that the Internet architecture has enabled for the past 35 years.

Net neutrality advocates often incorrectly insist, in this ex-parte and in other statements, that DiffServ is used nowhere on the Internet today, when in fact it is well known that it's used to implement a number of edge and transit services, as filings by network operators such as AT&T have indicated. It's certainly clear that commercial Internet services are routinely sold with Service Level Agreements which specify differentiated pricing for differentiated transport services.

Networks employ several means for identifying the desired class of service for particular packets above and below the level of the DiffServ option in the Internet Protocol; Quality of Service designators are found in IEEE 802.1d (VLAN,) 802.3 (Ethernet,) and 802.11e (Wi-Fi.) They are also found in DOCSIS, DSL, in BGP Community Attributes and in MPLS. The oft-repeated claim that all packets are equal on the Internet, or if they aren't, their inequality is strictly under user control, is false as a matter of architecture and of empirical fact, but no amount of evidence seems capable of putting the fiction to bed.

Even if there were no differentiation of packets on the Internet today, it would not follow that there would never be any legitimate reason to differentiate in the future. The architects of DiffServ anticipated "heterogeneous application requirements" driven not so much by the nature of networks as by the nature of the senders and receivers of information, machines, and the human sense organs. The requirement of the ear for voice communications with no more than 150 milliseconds of latency per packet is independent of network technology, as is the desire of file transfer programs to complete their work as quickly as possible. But the ear's requirement is an absolute feature of our biology, while the file transfer program's desire for timely completion is simply a machine preference. These activities don't have equal significance and they need not have equal priority as a general matter.

3. Openness and Neutrality are Contradictory Goals

The traffic management problem only becomes more serious as applications become more heterogeneous, but the insistence on absolute openness and absolute neutrality limit the ability of networks to perform the functions that users require of them. While "openness" and "neutrality" are both laudable goals that are urged on the Commission by filers such as the Open Internet Coalition, they are incompatible in some important ways.[9] A completely open Internet is one in which any form of content, any application, and any service is carried to the user's satisfaction, and a completely neutral Internet is one in which all packets move without favor. Openness, understood in this way as a commitment to satisfactory performance for heterogeneous applications, is different from and more important than "neutrality." This is not a novel or idiosyncratic observation; it was voiced by Tim Wu in the paper that introduced the term "network neutrality," to wit:[10]

> Proponents of open access have generally overlooked the fact that, to the extent an open access rule inhibits vertical relationships, it can help maintain the Internet's greatest deviation from network neutrality. That deviation is favoritism of data applications, as a class, over latency-sensitive applications involving voice or video. There is also reason to believe that open access alone can be an insufficient remedy for many of the likely instances of network discrimination.

While Wu places a higher value on neutrality, openness has more significance to the network user. The fact that the Commission has dropped the "neutrality" language in favor of the "openness" formulation in its current proceedings is an encouraging sign that

[9] "The court's decision makes uncertain the Commission's ability to implement several important parts of the National Broadband Plan, and threatens the ability of the Commission to adopt rules to protect an open and neutral Internet." Markham Erickson, "Comments of the Open Internet Coalition In the matter of Framework for Broadband Internet Service," July 15, 2010.

[10] Tim Wu, "Network Neutrality, Broadband Discrimination," *SSRN Electronic Journal* (2003), http://www.ssrn.com/abstract=388863.

indicates how far the debate has progressed. "Internet Openness" comprehends the prohibitions on operator blocking and degrading that are the constructive elements of the network neutrality agenda without imposing the tacit ban on operator actions to prevent harm from the cross-application degradation effects that are a much more concrete and immediate concern to Internet users than the conjectural problems are today.

Genuine openness can only be achieved on the network of heterogeneous applications when applications can voluntarily communicate their individual requirements to networks, and networks can differentiate services – and allocate costs – in a manner appropriate to application requirements. This is because the greatest barrier to the success of any given application or other traffic flow on the Internet is the simultaneous behavior of other applications on shared network components. If networks are merely passive conduits, latency-insensitive high-traffic applications will degrade the performance of latency-sensitive applications. This is the case when unmanaged P2P file transfer flows compete for bandwidth with latency-sensitive VoIP, and it's not simply alleviated by adding bandwidth in edge networks.

4. There are Limits to Bandwidth Growth

There are practical limits to the ability of network operators to solve congestion problems by adding bandwidth in any case because of the shared nature of the Internet, the limited number of Internet Exchange (IX) points, the limited bandwidth of Ethernet switches employed at the IX's, and the limited number of BGP routes that modern routers can process at "wire speed." Internet packets are aggregated from slower to faster links as they move from the home or office to the Internet Exchange, typically in two stages, first at the end of the first mile (in a telephone company Central Office (CO) or a cable system Cable Modem Termination System (CMTS)). They are aggregated at least once again on their way to the Internet Exchange in some form of Concentrator (a network switch or router.) The highest speed that's commonly supported at the IX is 10 Gbps. If the links between home and CO are 1 Gbps, for example, no more than ten (times the oversubscription factor) homes could be aggregated onto a single Concentrator. The number of Concentrators would therefore grow, and as they did, the number of BGP routes would increase by orders of magnitude above the 300,000 that are in use today, thereby raising the cost of Internet connectivity beyond the reach of the average consumer.[11] Consequently, consumer bandwidth can only increase in modest proportion to the most commonly employed high-speed Ethernet switches at IX's, which are limited by the state of technology at any given time. When the bandwidth demands of popular applications require faster rates of increase in bandwidth, they must be moderated by traffic management.

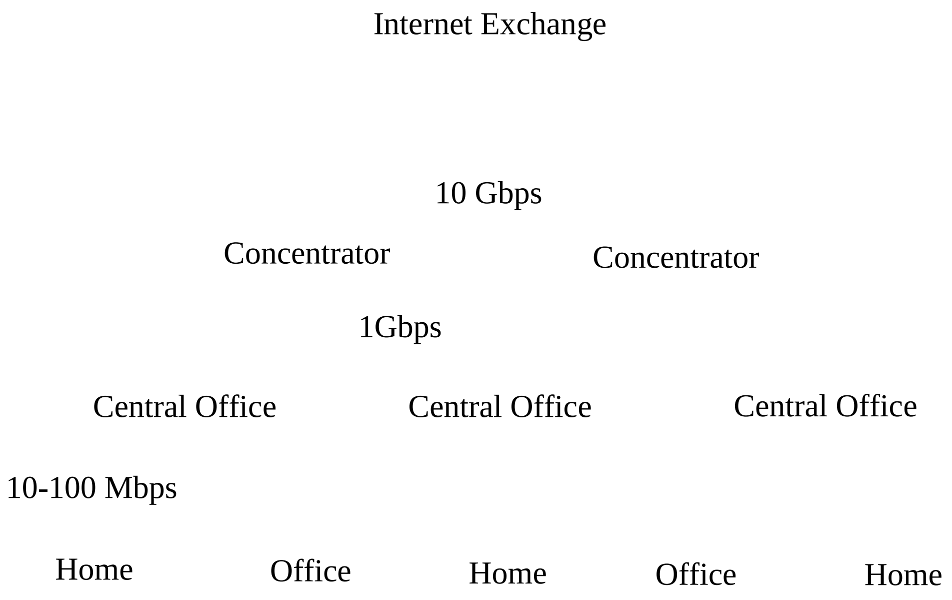

Figure 1: Internet aggregation hierarchy

To be effective, bandwidth increases have to be made across the entire Internet and not just in certain parts of the "last mile." Moderating bandwidth growth therefore helps keep Internet connectivity affordable and prevents the buildup of bottlenecks when some portions are sped up more than others. The Internet is a dynamic system, where users of the fastest links enjoy exponentially better service than the users of the slowest links. This fact is not well understood by the advocates of strict regulations on traffic

[11] If the Internet user population were to decrease due to rising costs, price reductions brought about by the next generation of technology would win at least some of them back.

management as they tend to approach the network economics from a narrower point of view. The TCP sliding window has non-neutral effects on network traffic because faster paths can open their congestion window faster than users of slower links can. Therefore, faster link users inevitably take shared capacity away from users of slower links; TCP congestion control may be first-come first served, but faster users are served before slower ones.

5. Content Delivery Networks Prioritize Packets

One of the arguments that has been made repeatedly by network neutrality advocates is that edge caching Content Delivery Networks (CDNs) do not have the same "discriminatory" effects that packet prioritization has. One filer has asserted:[12]

> First, last-mile broadband access providers are uniquely positioned as a technical matter vis-a-vis all other entities connected to the Internet. Last-mile broadband access providers' networks act as the on/off ramps for Internet traffic, so that every packet of Internet traffic must traverse the networks and devices under their control. Because they own and operate physical last-mile networks, including the routers closest to the end user customer, broadband providers are able to inspect, act upon, and apportion capacity for all online traHic - including third party Internet traffic -- that traverses their networks. By contrast, other entities on the Internet, including applications and content providers, can view and interact with only their own traffic.

While last mile networks are uniquely positioned as a matter of fact, this claim naively conflates "inspection" with "interaction." Every network operator – whether a last mile provider or a transit provider – can inspect packets, just as every web indexer can inspect content. But the ability to inspect is neither a precondition to the ability to interact with nor to the ability to affect traffic on shared facilities. An accident on a highway causes congestion that affects traffic far beyond the vision of the parties involved in the accident, after all. Content delivery networks increase the speed at which applications can deliver packets to nearby last mile networks in exactly the same way that paid peering and other forms of premium service do.

[12] Vijay Gill, "Declaration of Vijay Gill," April 26, 2010, http://fjallfoss.fcc.gov/ecfs/document/view?id=7020438892.

There are two technical reasons for this:

1. Delays in the delivery of packets are imposed by each step (or “hop”) in the routing process of packets across the Internet. The smaller the number of hops, the greater the opportunities for delay. The value proposition of the CDN is predicated on their ability to accelerate delivery, which they do by eliminating hops.
2. The short path from the CDN to the last mile gives CDN traffic an advantage is terms of the contention for bandwidth that is mediated by the TCP congestion avoidance mechanism. Networks react to congestion by dropping packets, and TCP reacts to packet loss by reducing its “window size” and then increasing it as packets are successfully acknowledged. Acknowledgements flow more rapidly on short paths than on long paths, so CDNs are allowed by convention to transmit more packets across congested links than faraway sources are. All other things being equal, the throughput of a TCP sender is inversely proportional to the delay between source and destination. CDNs exploit this feature of Internet congestion avoidance to effectively increase the priority that their users enjoy on any given last-mile network.
3. Effectively, the CDN moves its packets to the head of the Internet’s system-wide queue for a given destination network. This is not to say that there is a single system-wide *physical* queue so much as to say that there is in effect a system-wide *system* of physical queues that describes packet delay in total. Consider the following diagram:

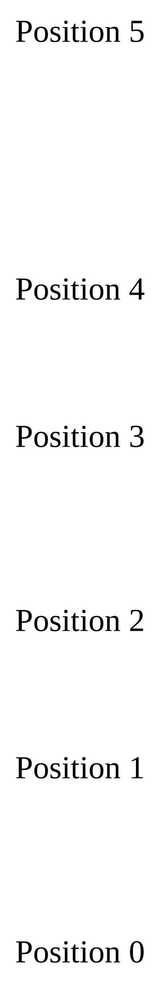

Figure 2: Content Delivery Network Prioritization

At every stage in the transfer of a packet from the faraway source to the last mile network, the packet incurs a queuing delay that is absent from the path between the CDN and the last mile network. The queuing delay is in part a result of the latency of packets moving across the cable from one router to another, and also from the delay caused by packets that reached the router first.

The cumulative effect of these queuing delays to decrease the share of last-mile bandwidth available to the sender in relation to the share that will be obtained by the nearby CDN, which comes about in part by increasing the delay the sender undergoes in expanding his TCP window. Moving packets to the head of the router queues in each router – as the operator may do for premium traffic – only has the effect of reducing the discrepancy between far away traffic and local CDN traffic.

In non-content oriented applications such as video conferencing, jumping to the head of each router queue is the only form of acceleration that's possible, but the effect produced on queued but non-prioritized traffic is no greater than the delay caused by the CDN. In most cases, it will be less because CDNs tend to deal in high bandwidth applications such as high definition video streaming while communication applications tend to have more modest requirements.

In this example, the CDN competes fairly with packets coming from the Router at Position 0, but has an advantage over packets from other routers and from the faraway sender by virtue of its superior ability to replace both transmitted packets and dropped packets with new ones. This is why CDNs are successful businesses.

On shared medium packet switched networks, every packet affects every other packet in the same path, independent of the observation point. As more bandwidth is consumed by one party, less is available to others.

There is an interesting economic consequence of both CDNs and premium delivery services in that they both have the effect of making more money available to network operators for capacity upgrades: The CDN reduces operator transit costs by reducing the number of hops over which the operator must carry the packet, and the premium service provides a direct subsidy to upgrade overall capacity. As the premium service is not in use all the time, non-premium users benefit from the upgrades to shared facilities. Both CDNs and premium services produce good effects.

6. There Are Other Ways to Obtain Authority

The FCC's quest for authority to regulate the Internet's edge networks appears to be predicated not only in the desire to impose an Open Internet policy framework on edge providers, but on certain elements of the older and significantly different net neutrality agenda. A truly open Internet – one in which new applications run successfully side-by-side with traditional ones – is not neutral, it's fair and active. It's a system in which applications communicate their requirements to the network and the network communicates its conditions to applications, where every application has a basic level of service and premium service is available on demand on a non-discriminatory basis. The Commission seems to be overly concerned with its legal authority at the moment, and insufficiently attentive to the pressing policy issues. We have different priorities, as we would prefer that sound policy precede enforcement power. It is also our belief that the correct policy can largely be implemented through a voluntary system of compliance, as we shall explain shortly.

Ultimately, the clearest path to such regulatory power is for the Congress to explicitly grant it to the FCC, and this will most likely happen when the current movement in Congress to clarify the FCC's Internet regulation power is concluded. In the meanwhile, the FCC can employ voluntary means to protect the Internet from interim danger, however remote that danger might be. The preferred approach would be for the agency to request voluntary compliance by the relevant parties with a consensus interim policy framework. There is general agreement that the open Internet should enjoy meaningful protection from bad actors, even if there is not a consensus that principles of neutral management or non-management advance that goal. A truly open Internet is free of arbitrary blocking and degrading by both operators and applications; a neutral Internet is one in which high appetite applications on high speed links take bandwidth away from others. Openness is not advanced by neutrality, but only by active management.

Therefore, we recommend that the FCC should convene an expert advisory panel of Internet firms, academics, think tank experts, and public interest lobbyists to draft and to enforce guidelines that represent the degree of consensus that exists today. To the extent that the FCC finds it necessary to go beyond the consensus to enact specific protections, it can do so, but at the expense of weakening its enforcement power by virtue of working with a smaller number of voluntary participants.

The regulatory framework could then be enforced by the mandated refusal of consenting Internet firms to do business with non-compliant dissenters. This power would effectively isolate non-compliers from the Internet, which would spell doom to their economic prospects and serve as a powerful motivator to comply with consensus guidelines.

The implementation of isolation actions is a serious and powerful remedy that must be protected by due process and clearly lawful. Therefore, the Commission and the voluntary body will need to develop clear guidelines on the process of enforcing isolation, such as requirements for a super-majority vote, consent of the FCC, recourse for the affected party, and freedom from operator liability under existing contracts.

The consensus guidelines would remain in effect until Congressional action clarifies the FCC’s role and legitimate policy framework is developed that carries the force of law.

Disputes over compliance would be mediated by the advisory panel after careful review of the facts. This is a workable system that resembles advisory and certification processes that are already employed in network businesses such as the Wi-Fi Alliance, but with greater enforcement power. It is also practiced to some extent by Internet operators today, who blacklist known distributors of malware and networks that host criminal enterprises.

7. Conclusion

The immediate desire to obtain legal authority over the Internet is motivated by the ambition to impose Open Internet conditions on edge network operators. Open Internet conditions are generally good, but they conflict with network neutrality goals. Proponents of network neutrality fail to recognize the difference, and propose a system of regulation that stands in conflict with the clearly stated goals and purposes of important Internet standards such as RFC 2475, “to permit differentiated pricing of Internet service” in order “to accommodate heterogeneous application requirements.”[13] Network neutrality is therefore in conflict with the Internet’s design principles.

Pending clarification of its role as an Internet regulator by the Congress, the FCC should convene a panel of Internet stakeholders to develop interim guidelines consistent with RFC 2475 and the rest of the Internet’s standards and best practices. With sufficiently broad membership, this stakeholder organization should have the necessary power to ensure that the Internet remains an open platform for innovation and consumer choice pending Congressional action. The Internet is sufficiently healthy and vibrant that there is no reason to believe an interim system of self-governance cannot be effective. Rather, the Internet’s history suggests that self-governance is the most effective long-term mechanism for ensuring both progress and openness.

[13] Blake et al., “RFC 2475 - An Architecture for Differentiated Services.”